Curso: Introducción a los Métodos Numéricos en Astronomía
Prof.: Cecilia Mateu J.

Fecha límite de entrega: martes 26 de marzo 2013

Tarea 6 - Inferencia Bayesiana

La entrega de la tarea debe hacerse por correo electrónico a cecilia.mateu@gmail.com. La tarea debe consistir en un archivo llamado inicialapellido_nombre_tareaN.tar.gz o .zip que contenga:

- Los códigos .py correspondientes a cada problema de la tarea
- Los gráficos (en formato .pdf, .png o .jpg) y/o tablas que se soliciten en la tarea,
- Un archivo de texto (texto simple, .doc o .odt) con las respuestas escritas que se soliciten en la tarea
- Un archivo README (en texto simple preferiblemente) que indique cómo se debe correr cada código desde la línea de comando.

Recuerde que para ser evaluado favorablemente el profesor ***debe*** poder correr los códigos y reproducir los resultados, figuras y/o tablas entregados. No será aceptable entregar o hacer referencia a resultados sin los códigos correspondientes que los generen.

Inferencia Bayesiana

1. En el siguiente problema considere que Ud. tiene un conjunto de N medidas m_i de la magnitud aparente en la banda V para una estrella del Red Clump. Éstos datos se suministran en el archivo ***datos_rc1.dat*** que contiene magnitud aparente y su correspondiente error en la primera y segunda columna respectivamente. El objetivo es determinar la distancia D a la estrella (y su distribución posterior de probabilidad) utilizando inferencia bayesiana.

   a. Suponga que las magnitudes medidas m_i tienen errores gaussianos, con σ_i dado por el error reportado para cada medida en el archivo ***datos_rc1.dat***.

   b. Escriba la distribución posterior de probabilidad $p(D|\{m_i\})$ de la distancia D, dados los datos observados $\{m_i\}$. Suponga un prior uniforme en D (con unos límites que le parezcan razonables). Recuerde que la magnitud aparente m de una estrella se relaciona con su distancia D (en pc) a través de la Ec. de Pogson

$$m - M_V = 5 \log D - 5$$

   suponiendo que no hay extinción. Asuma que la magnitud absoluta de una estrella del Red Clump es M_V=+0.55. **(4 pt)**

c. Escriba un programa que lea los datos del archivo ***datos_rc1.dat***. Calcule y grafique la distribución posterior $p(D|\{m_i\})$ derivada en la parte b. **(2 pt)**

d. Demuestre ***analíticamente*** que el mejor estimado D_o de la distancia se puede despejar a partir de la relación

$$5 \log D_o - 5 = \bar{m} - M_V - \frac{\ln 10}{5w}$$

donde $\bar{m}$ es el promedio pesado dado por

$$\bar{m} = \left(\frac{1}{w}\right) \sum_{i=1}^{N} \frac{m_i}{\sigma_i^2}$$

donde w es la suma del inverso de los errores cuadráticos dada por

$$w = \sum_{i=1}^{N} \frac{1}{\sigma_i^2}$$

(5 pts)

e. Grafique el valor obtenido de D_o como una línea vertical en el gráfico de la posterior que hizo en la parte c. Este estimado de Do, ¿debe coincidir con el máximo de la posterior que Ud. calculó numéricamente? ¿por qué? **(2 pts)**

f. Escriba la expresión analítica de la distancia D_{mal} que Ud. hubiera hecho si no hubiera usado la inferencia bayesiana (i.e. à la laboratorio 1) **(1 pts)**. Discuta cómo se compara este resultado con el valor de D_o obtenido en la parte anterior. Discuta en qué se diferencian ambos cálculos y qué implicaciones tiene el resultado de la parte anterior. **(3 pts)**

g. Discuta qué espera que ocurra con la posterior $p(D|\{m_i\})$ y con los valores de D_o y D_{mal} a medida que los errores σ_i sean cada vez menores. **(3 pts)**

FIN